Economics 142: Behavioral Economics **Spring 2008**
Vincent Crawford (with very large debts to Colin Camerer of Caltech, David Laibson of Harvard, and especially Botond Koszegi and Matthew Rabin of UC Berkeley)

Introduction to Behavioral Economics
Like the subject matter of behavioral economics, this course is divided into two parts:

The first half studies "behavioral decision theory" models of individual decision making.

● Standard economic models of decision making rest on useful but extreme assumptions about people's preferences, how they form their beliefs and adjust them in response to new information, and how much will-power and self-control they have in pursuing their goals.

● Behavioral decision theory studies the consequences of modifying those assumptions in psychologically plausible directions suggested by introspection, experiments, and sometimes field observation. (These are summarized shortly and discussed in detail later.)

● The premise is that there are systematic and economically important deviations from standard assumptions; otherwise we would be happy to stick with the standard model.

● The modified assumptions are then used to re-do some standard economic analyses of decision making, with the goals of explaining empirical puzzles and giving us ways to think about economic issues that are not well handled by standard models.

The second half of the course studies "behavioral game theory" models of how people make decisions in interactive situations ("games") in which a person's own outcome and welfare are influenced by others' decisions as well as her/his own decisions.

● Standard economic models of decision making in games rest on useful but extreme assumptions about how people predict others' decisions.

● Behavioral game theory studies the consequences of modifying those assumptions in plausible directions suggested by introspection and/or experiments.

● The premise is that there are systematic and economically important deviations from standard assumptions; otherwise we would be happy to stick with the standard model.

● The modified assumptions are then used to "re-do" parts of standard economic analyses of decision making in games, again with the goal of explaining empirical puzzles and giving us ways to think about economics issues that are not well handled by standard models.

Even these vague descriptions make it clear that behavioral economics is not a "field" like international trade or industrial organization: It is distinguished by its methods, not by its subject matter.

They also suggest that behavioral economics is, first and foremost, economics: just a somewhat different way of doing it.

Most people who do behavioral economics hope that it will eventually merge with and enrich mainstream economics, hence losing its own identity as a field.

This is what happened with econometrics, and more recently with traditional game theory.

This point of view is very nicely expressed (in unpublished notes) by Matthew Rabin:

Premise 1: Most (not all) facts about people that seem to be true and seem to be economically relevant–even if these assumptions didn't appear in the prior generation of economics textbooks are ... both true and economically relevant.

Premise 2: "Untraditional" or unfamiliar assumptions, including those that imply various limits to rational utility-maximization, can and should mostly be studied using exactly the same set of tools and approaches economists are used to (i.e., formal mathematical models and statistical tests using laboratory and especially field data), using exactly the same scientific criteria (good predictions, parsimony, etc.) as economists are used to. The sole difference in methods and goals of most economists is the broader array of aspects of human nature we study.

Premise 3: Not only are familiar economic methods great, but to a very large extent so are familiar economic assumptions. The fact that there are limits to the correctness and applicability of these assumptions does not mean that they aren't often exactly the appropriate assumptions–nor that they aren't tremendously useful even when not exactly right. The material in this course is not meant as a replacement of, but as an enhancement of–and eventual component of–mainstream economics…. The eventual goal … is that … "behavioral economics" will eventually disappear as a separate or isolated field….

Introduction to Behavioral Decision Theory

Before we start introducing behavioral decision theory, I ask you to take a few minutes to fill out a short survey that will help us think about its assumptions later on.

Your answers will not be graded, or even evaluated; and you are not even asked to put your name on the form.

But the results will be more interesting and useful if you answer the questions carefully.

Focusing for now on individual decisions, what does mainstream (standard neoclassical) economics normally assume?

Homo Economicus (possibly unlike Homo Sapiens):

● Is perfectly rational, making choices that consistently maximize some exogenous, stable set of preferences that depend on absolute levels of outcomes (rather than changes), even with uncertainty (in which case preferences are expected utility), and even in dynamic situations (in which case preferences are discounted sums of per-period expected utilities)

● Is also perfectly rational in the sense of costlessly and correctly making logical nonprobabilistic inferences and applying the laws of probability to process information and make probabilistic judgments (Bayes' Rule, contingent reasoning, option value)

● Has perfect will-power and the ability to make and follow intertemporal plans (even contingent ones), with no conflict between the preferences of current and future selves

● Is almost always also assumed to be perfectly self-interested, caring exclusively about her/his own consumption, though this assumption is not essential to mainstream theory

Familiar as they are, nothing essential depends on whether these assumptions are considered "standard".

Even rationality, though very important to the way economics is done, is not an essential axiom without which no coherent theory is possible.

What are "standard" assumptions is only a convention of the discipline, subject to change when different assumptions appear to be more useful.

In many settings the standard assumptions are reasonable stylizations of the "facts" that most of the people whose behavior we wish to analyze are usually self-interested and well-informed, with coherent goals, and reasonable skill at making plans to realize them.

The standard assumptions also embody a kind of methodological humility, preventing us from assuming we know more about people's goals and possibilities than they do.

But in other, equally important settings, standard assumptions are unreasonable as descriptions of behavior.

This means more than that they are literally incorrect; all behavioral assumptions are incorrect to some degree. It means that they are systematically incorrect, and that the incorrectness has economically important consequences; otherwise we would be happy to stick with the standard model.

In such settings the standards assumptions have sometimes significantly hindered research.

They have allowed/encouraged economists to ignore research that directly explores preferences, beliefs, information processing, and other determinants of human behavior.

There was no need to study the structure of preferences directly, because under the standard assumption that choices maximize preferences, preferences can be inferred from choices via revealed preference.

There was no need to study belief formation or information processing, because it was assumed to be completely determined by rationality postulates.

Behavioral economists, by contrast, have to be open to alternative ideas and evidence about preferences, beliefs, and information processing.

In such settings, behavioral economists (unlike some mainstream economists) are willing to consider deviations from standard assumptions in directions suggested by behavioral evidence, if it yields better explanations of outcomes than standard models do.

I now give some illustrative examples of systematic and important deviations from standard assumptions, starting with your survey responses and then discussing analogous phenomena in the "field" (the economy outside the classroom or laboratory).

The examples are meant only to make the abstract statements above more concrete and convince you that there is something worth further study here; we'll discuss these and related issues more carefully later in the course.

I divide the examples into three groups, each of which defines a direction or directions in which the standard model can be improved by making more realistic assumptions:

• Choice under uncertainty (or certainty)

• Probabilistic judgment

• Intertemporal choice

These groups are the main components of behavioral decision theory studied in the course.

(For lack of time, I omit a fourth group, social preferences, which includes deviations from purely self-interested preferences such as altruism, envy, spite, and reciprocity. This group is also interesting, but well-covered in Economics 141, Experimental Economics.)

Choice under uncertainty

First consider the answers to questions 1 (a and b) of the survey.

(Actually there were two versions of each question but 3, one answered by half of you, one by the other half. As we will see, comparing the answers across the different versions yields more information than examining the answers in isolation. Because the different versions were assigned (essentially) randomly, and the class is fairly large, it's reasonable to assume that the differences in answers across version reflect something systematic about people in general, rather than differences across the groups that answered each version.)

1a. Would you choose to lose $500 for sure or to lose $1000 with probability 0.5?

1b. Would you choose to receive $500 for sure or to receive $1000 with probability 0.5?

● Most people who answer questions like 1a choose to lose $1000 with probability 0.5 rather than losing $500 for sure, suggesting "risk-loving" behavior with respect to losses. (This suggests that people dislike losses so much they are willing to take a fairly large, equal-expected-money-outcome risk just to reduce the probability of a loss.)

● By contrast, most people who answer questions like 1b chose to receive $500 for sure rather than $1000 with probability 0.5, suggesting "risk-averse" behavior with respect to gains.

These responses differ from standard assumptions in two ways:

● Standard theory tends to assume that people are risk-averse for all choices, unlike in the risk-loving choices with regard to losses

● More importantly, standard theory almost always assumes that preferences are defined over absolute, final levels of outcomes (or probability distributions of them) rather than changes (or distributions of them), which strongly suggests that choices should be qualitatively the same—both risk-loving or both risk-averting—for 1a and 1b.

Why only "strongly suggests"?

If the people who answered 1a are as wealthy as those who answered 1b on average, then the gains in 1b versus the losses in 1a tend to make the latter a bit richer than the former.

In principle, this could make them sufficiently more risk-averse to flip the responses the way they usually flip.

But this is not a plausible explanation, because the difference is not large enough to explain such a big flip.

(And more refined questions/experiments make it look even less plausible.)

We will see a lot of evidence that the most plausible explanation of the flip is what behavioral economists call "reference-dependent" preferences: preferences defined, not over absolute amounts, but over gains or losses defined relative to some "reference point."

In defining reference-dependent preferences to explain the flip, it is natural to take the reference point as the status quo before the choice, but we will have to think harder about how to define it in other applications.

We will also have to think harder about "mental accounting" and "bracketing": how the individual groups choices and risks together in thinking about them.

Later on we will see several important examples of using reference-dependent preferences to explain economic puzzles. To name two, reference-dependent preferences can explain:

● The phenomenon that race-track bettors tend to bet more on long shots near the end of the day at the track. (What's the reference point? What's the bracket?)

● The phenomenon that house sellers who paid more for their houses (a sunk cost in standard theory) set asking prices that are higher, controlling for quality, so that they tend to take longer to find a buyer, but to sell at a higher price. (What's the reference point? What's the bracket?)

Now consider the answers to questions 2 (a and b).

2a. Would you choose to receive $3,000 for sure or to receive $4,000 with probability 0.8?

2b. Would you choose to receive $3,000 with probability 0.25 or $4,000 with probability 0.2?

● Most people who answer questions like 2a prefer $3,000 for sure.

● But most people who answer questions like 2b prefer $4,000 with probability 0.2.

This by itself is not clear evidence that something other than distributions of final outcomes matters. But re-frame it as a two-stage decision as follows:

In the first stage, with probability 0.75 the process ends with you winning $0, and with probability 0.25 you move into the second stage. In the second stage, you choose between receiving $4,000 with probability 0.8 and $3,000 for sure. (Your choice here must be made before the outcome of the first stage is known.)

This is mathematically identical to the choice in 2b, but here, unlike in the equivalent choice in 2b, most people choose $3,000 for sure in the second stage. Once they see the chance of getting $3,000 for sure, they think about that risk very differently: not only the distribution of final outcomes matters. (This question also relates to intertemporal choice.)

Probabilistic judgment

Now consider the answers to question 3 (which was the same for all).

3. Suppose that one out of a hundred people in the population have HIV. There is a test for HIV that is 99% accurate. This means that if a person has HIV, the test returns a positive result with 99% probability; and if a person does not have HIV, it returns a negative result with 99% probability. If a person's HIV test comes back positive (and you know nothing else about her/him), what is the probability that s/he has HIV?

Most people answer 99%. This is wrong. Perhaps the reasoning went as follows:
● An HIV-negative person will probably receive a negative result (99% chance)
● An HIV-positive person will probably receive a positive result (99% chance)
● Conversely, if a person tested positive, she is likely to be HIV-positive (99% chance)

The problem with this is that it ignores the base rate ("one out of a hundred people in the population have HIV"); ignoring the base rate makes you systematically overestimate the probability of rare events and underestimate the probability of common events. Taking the base rate into account requires at least an intuitive understanding of Bayes' Rule: An HIV-negative person is 99 times less likely to test positive than an HIV-positive person, but there are 99 times more HIV-negative people. These cancel out, so the probability that a person testing positive has HIV is exactly 50%. (We'll see how to do the algebra later)

To see this point in a different way, consider the answers to questions 4 (a and b).

4a. Jack's been drawn from a population which is 30% engineers and 70% lawyers. Jack wears a pocket protector. Use your own estimate of the respective probabilities that engineers and lawyers wear pocket protectors to estimate the probability that Jack is an engineer.

4b. Jack's been drawn from a population which is 30% lawyers and 70% engineers. Jack wears a pocket protector. Use your own estimate of the respective probabilities that lawyers and engineers wear pocket protectors to estimate the probability that Jack is an engineer.

If p_1 is the probability that Jack is an engineer in 4a (30% engineers, 70% lawyers) and p_2 is the probability that Jack is an engineer in 4b (30% lawyers, 70% engineers), then using Bayes' Rule we can show that (independent of the probabilities that engineers and lawyers wear pocket protectors, which cancel out) $[p_1/(1- p_1)]/ [p_2/(1- p_2)] = (3/7)^2 \approx 18\%$. But people's average estimates for p_1 and p_2 are virtually the same.

Again, people ignore the base rate and systematically overestimate the probability of the relatively rare event and underestimate the probability of common events.

Later on we will see several important examples of using this and other biases in probabilistic judgment to explain economic puzzles.

Intertemporal choice

Finally, consider the answers to questions 5 (a and b).

5a. Suppose I could give you either $100 in cash right now or $x in cash in two weeks. What is the x for which you would be indifferent between the two options?

5b. Suppose I could give you either $100 in cash in 52 weeks or $x in cash in 54 weeks. What is the x for which you would be indifferent between the two options?

People typically answer that x = about 138 in 5a but x = about107 in 5b. Thus they are much more impatient about a two-week delay of gratification when immediate gratification is possible than when the delay starts only a year from now.

This is inconsistent with standard assumptions about intertemporal preferences, where the trade-off between consumption now versus two weeks from now is not systematically different from the trade-off between consumption in 52 weeks and in 54 weeks.

As we will see, this "present bias" in intertemporal preferences helps to understand some important economic puzzles that resist explanation under standard assumptions.

To name one, in the US in 2001, there were one billion credit cards (5 per adult?), people with credit card debt averaged almost $6000 in balances on which they paid interest, on which they paid more than 18% interest on average, making interest payments of about $1000 per year per person carrying debt.